

Routage en boucle ouverte dans deux files $G/M/1$ en parallèle

Bruno Gaujal and Emmanuel Hyon and Alain Jean-Marie

INRIA/LIP ENS Lyon, 46 Allée d'Italie, F-69364 Lyon, France. Bruno.Gaujal@ens-lyon.fr

LORIA-INRIA, UMR 7503, 615 route du jardin botanique, F-54606 Villers-les-Nancy, France. hyon@loria.fr

LIRMM, Université Sciences et Techniques du Languedoc, 161 Rue Ada, F-34392 Montpellier, France. ajm@lirmm.fr

Dans cet article nous nous intéressons au routage en *Boucle ouverte* dans deux files d'attente en parallèle lorsque les services et les inter arrivées avant routage sont exponentielles. Notre but est de trouver la politique de routage optimale au sens où elle va minimiser le temps d'attente moyen. Pour ce faire nous étudions une file d'attente dont le processus d'entrée est un processus de Poisson échantillonné.

Keywords: Filtrage d'un processus de Poisson, Routage en boucle ouverte, Méthode du noyau, Mot de Sturm.

1 Introduction

Les nouveaux besoins, que ce soit en terme de Qualité de Service ou bien en faisabilité de contraintes temps réel, qui apparaissent dans les réseaux de télécommunications rendent l'apparition de nouvelles politiques de routage et de contrôle d'admission nécessaires, l'augmentation des débits ne permettant pas de faire face à tous les problèmes. De plus cette augmentation des débits conduit à des situations où le traitement de l'information disponible ne peut être effectué avant le routage. Ceci se produit notamment dans les réseaux à fibres optiques.

C'est pourquoi nous nous intéressons dans ce travail au routage en *boucle ouverte*. Cela signifie que le contrôleur ne dispose comme information que de la suite de ses actions antérieures ainsi que des paramètres du système et ne connaît jamais l'état global du réseau. On peut se référer à [Alt00] ainsi qu'aux références incluses pour avoir un aperçu des multiples travaux sur le routage optimal dans des réseaux.

Nous nous intéressons à un système composé de deux files d'attente en parallèle. Le routage consistant alors en un mot m binaire infini dans lequel chaque lettre représente l'affectation du client dans l'une des files. Lorsque les temps de services et les inter arrivées sont déterministes le calcul du routage optimal dans deux files a été fait dans [GH01] tandis que [GH02] traite du routage optimal entre plusieurs circuits qui trouve une application par exemple dans MPLS-OMP ([Vil99]).

En vue de généraliser ces travaux, nous nous intéressons au cas où les inter arrivées avant routage et les services suivent une loi exponentielle. Ceci nous conduit à l'étude d'une file d'attente $G/M/1$ dont le processus d'entrée est une sélection déterministe des clients qui arrivent suivant un processus de Poisson. Notre problème peut être modélisé par processus de Markov en deux dimensions en réalité un processus de quasi naissance et mort qui peut être vu comme une généralisation du processus de naissance et mort de la file $M/M/1$.

Notre manière de calculer de la probabilité stationnaire du processus de quasi naissance et mort présentée ici diffère des méthodes classiques. En effet notre méthode dite *méthode du noyau* consiste à calculer les solutions d'une équation fonctionnelle vérifiée par la fonction génératrice de la probabilité stationnaire de cette chaîne de Markov. Si certains aspects probabilistes sont perdus, par contre la majorité des calculs se ramène à l'approximation de zéros de fonctions polynomiales et à l'inversion de matrices creuses.

Notre papier s'organise comme suit dans la section 2 nous explicitons plus précisément le routage échantillonnage et construisons le processus d'entrée échantillonné dans un cadre général. Nous présentons dans la section 3 la méthode du noyau et calculons le temps d'attente moyen dans une file. La section 4 traite quant à elle du calcul de la politique de routage optimale : optimale au sens où elle minimise le temps d'attente moyen des clients. Enfin dans la section 5 est consacrée aux applications numériques.

2 Routage et construction du processus d'entrée

Nous allons préciser dans cette partie la manière dont le routage et par la même l'échantillonnage fonctionnent. On se place pour l'instant dans un cadre général dans lequel le processus des arrivées dans le système est un processus ponctuel stationnaire ergodique d'intensité $1/\lambda$ noté \mathcal{X} . On s'intéresse à un système composé d'un routeur et de deux files en parallèles $G/G/1/\infty$ FIFO notées Q^1 et Q^2 dont les services sont ici aussi des processus stationnaires ergodiques d'intensités respectives $1/\mu^1$ et $1/\mu^2$ indépendants entre eux et indépendants des inter arrivées.

Le routage fonctionne de la manière suivante on se donne un mot binaire m infini. Si $m(n)$ la $n^{\text{ième}}$ lettre de m vaut 1 ceci indique que le $n^{\text{ième}}$ client (ou paquet) est routé dans la file Q^1 sinon il est routé dans la file Q^2 .

Ainsi les processus d'entrée dans chaque file sont des filtres du processus d'entrée dans le système c'est à dire un processus échantillonné d'un processus stationnaire ergodique. Nous allons montrer un certain nombre de propriétés de ces processus filtrés.

Définition 1 (Processus filtré). On note par m^j le mot binaire décrivant les entrées dans la file Q^j . Si le $n^{\text{ième}}$ client admis dans le système est admis dans la file Q^j alors la $n^{\text{ième}}$ lettre de m^j ($m^j(n)$) vaut 1 et 0 sinon. On appelle $\mathcal{X}(m^j)$ le processus ponctuel filtré par m^j du processus \mathcal{X} .

Introduisons les notations suivantes.

Définition 2. Soit m un mot on note par $m_{[l]}$ le préfixe de longueur l du mot m , on note par a le nombre de 1 de $m_{[l]}$. On appelle pente du mot m la limite (quand elle existe) du taux a/l .

On appelle shift de m le mot infini périodique $S_k(m)$ tel que $S_k(m)(n) = m(n+k)$.

A partir de la définition suivante on note alors par $m_{[l]}^1$ et $m_{[l]}^2$ les préfixes de longueur l des mots m^1 et m^2 et par a^1 et a^2 le nombre de 1 de ces préfixes. On note par m_l la concaténation à l'infini de $m_{[l]}^1$ et $m_{[l]}^2$.

On considère alors Q comme une file $G/G/1$ dont le processus ponctuel d'entrée est $\mathcal{X}(m)$ et dont le processus d'entrée marqué (le processus dans lequel à chaque arrivée on associe un temps de service) est noté $\mathcal{X}(m, \sigma^j)$. Nous présentons ici un théorème qui justifie la suite du travail.

Théorème 3. Soit l un entier naturel fini. Soit U une variable aléatoire de loi uniforme sur $\{1..l\}$.

-i) Le processus $\mathcal{X}(S_U(m_l^j), \sigma^j)$ est un processus ponctuel marqué stationnaire ergodique.

-ii) Lorsque le processus d'entrée est $\mathcal{X}(S_U(m_l^j), \sigma^j)$, la file Q^j est stable si et seulement si

$$\frac{\lambda a^j}{l \mu^j} < 1. \quad (1)$$

-iii) Lorsque le processus d'entrée est $\mathcal{X}(m^j, \sigma^j)$ la file Q^j est stable si et seulement si

$$\limsup_{l \rightarrow \infty} \frac{\lambda a^j}{l \mu^j} < 1.$$

Démonstration. L'item i) résulte d'un court calcul de Palm tandis que ii) découle du schéma de Loynes. Pour plus de précisions voir [BB92]. \square

3 Calcul de la distribution stationnaire

On s'intéresse dans cette section au calcul de l'espérance du temps d'attente en régime stationnaire dans une file Q^j lorsque le processus d'entrée \mathcal{X} est un processus de Poisson. Le filtrage étant effectué en suivant m^j . On ne considère dans cette section que des mots m et ainsi m^j périodiques. Par abus de notation comme m respectivement m^j sont périodiques dans la suite on appelle m respectivement m^j leur plus petite période. On note par l la longueur de la plus petite période et par \hat{a} le nombre de 1 durant la plus petite période. Le processus échantillonné s'il est stationnaire ergodique (voir section 2) n'est alors plus un processus de Poisson les inter arrivées n'étant pas non plus indépendantes entre elles.

On peut modéliser le comportement de cette file $G/M/1/\infty$ FIFO par une chaîne de Markov en temps continu, X_t dont l'espace d'état est $\mathbb{N} \times \{1..l\}$. La première composante représente le nombre de clients dans le système à l'instant t tandis que la deuxième composante correspond à l'indicateur de la lettre courante de m^j .

Comme le mot m^j fini est répétée à l'infini et de plus que les lois des arrivées et des services ne varient pas en fonction du nombre de clients dans le système notre processus X est alors un processus de quasi naissance et mort dont le générateur s'écrit :

$$Q = \begin{bmatrix} C & A_0 & 0 & 0 & \dots \\ A_2 & A_1 & A_0 & 0 & \dots \\ 0 & A_2 & A_1 & A_0 & \dots \\ 0 & 0 & A_2 & A_1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

avec $A_0 = \lambda J^1$, $A_1 = \lambda J^0 - (\lambda + \mu)I$, $A_2 = \mu I$ et $C = \lambda J^0 - \lambda I$, où I , J^0 et J^1 sont des matrices carrées de dimension $l \times l$. La matrice I est la matrice identité. Les matrices J^0 et J^1 sont les matrices dont les coefficients se calculent par : $(J^0)_{i,k} = (1 - m^j(i))\delta_{i,k+1}$ et $(J^1)_{i,k} = m^j(i)\delta_{i,k+1}$. L'indice k étant pris modulo l et le terme δ représentant le symbole de Kronecker qui vaut 1 si $i = k$ et 0 sinon.

Soit π la probabilité invariante du processus X , cette probabilité vérifie alors (voir par exemple [Ros83])

$$\pi Q = 0. \quad (2)$$

On note par π_n le sous-vecteur de π de dimension l dont la i^{eme} coordonnée représente la probabilité stationnaire d'avoir n clients lorsque on se trouve en $m^j(i)$. Ainsi π_0 est la probabilité stationnaire de n'avoir aucun clients dans le système. On a $\pi = (\pi_0, \pi_1, \dots)$.

Classiquement au lieu de déterminer π directement nous allons calculer sa fonction génératrice.

Définition 4 (Fonction génératrice). La fonction génératrice est la fonction $\Pi(z)$ de \mathbb{R} vers \mathbb{R}^l définie par

$$\Pi(z) = \sum_{n=0}^{\infty} z^n \pi_n.$$

3.1 Méthode du noyau

Nous présentons maintenant la méthode du noyau celle-ci va nous permettre de calculer la fonction génératrice. Si cette méthode est souvent utilisée en automatique ou en combinatoire elle l'est relativement peu dans le domaine de l'évaluation de performance exceptée (à la connaissance des auteurs) dans [JM97]. Ainsi on détermine une équation fonctionnelle que vérifie la fonction génératrice, cette équation dépendant d'un paramètre. Le terme appelé noyau comporte un certain nombre de singularités. C'est ces singularités du noyau qui vont nous permettre de calculer le paramètre de l'équation fonctionnelle et ainsi la fonction génératrice.

Lemme 5. Soit $\rho^j = \lambda/\mu^j$. Soit $K(z)$, le noyau, la matrice $l \times l$ définie par

$$K(z) = \begin{bmatrix} 1 - (1 + \rho^j)z & \rho^j z^{1+m^j(1)} & & \\ & \ddots & \ddots & \\ & & 1 - (1 + \rho^j)z & \rho^j z^{1+m^j(l-1)} \\ \rho^j z^{1+m^j(l)} & & & 1 - (1 + \rho^j)z \end{bmatrix}.$$

La fonction génératrice vérifie l'équation fonctionnelle

$$\Pi(z)K(z) = \pi_0(1 - z). \quad (3)$$

Démonstration. Par l'équation de balance globale (2) on obtient la récurrence suivante

$$\begin{aligned} \pi_0 C + \pi_1 A_2 &= 0, \\ \pi_{n-1} A_0 + \pi_n A_1 + \pi_{n+1} A_2 &= 0, \quad \forall n \geq 1. \end{aligned}$$

En multipliant la seconde équation par z^{n+1} et en sommant on obtient

$$\Pi(z)(z^2 A_0 + z A_1 + A_2) = (1 - z)\mu^j \pi_0,$$

ce qui nous donne $K(z)$. □

Étudions maintenant les singularités de $K(z)$. Plus précisément nous sommes intéressées par les singularités de $K(z)$ qui se trouvent dans le disque unité puisque $\Pi(z)$ est une série entière dont le rayon de convergence est 1. Soit $\Delta(z)$ le déterminant de la matrice $K(z)$. Nous prétendons (voir section 3.3) que $\Delta(z)$ est un polynôme qui a l racines dans le disque unité. On a alors le théorème suivant :

Théorème 6. *Si z_i est la $i^{\text{ème}}$ racine dans le disque unité de $\Delta(z)$ et si v_i est le vecteur propre (de dimension l) associé à la valeur propre 0 de $K(z_i)$, alors π_0 est l'unique solution du système :*

$$\left\{ \begin{array}{ll} (1 - z_i)\pi_0 v_i = 0, & \forall i \in \{1, \dots, l\} \text{ t.q. } z_i \neq 1 \\ \pi_0 \mathbf{1} = 1 - \frac{a^j \rho^j}{l}, & \text{pour } z_i = 1 \end{array} \right\}. \quad (4)$$

où $\mathbf{1}$ est le vecteur dont les entrées sont toutes égales à 1.

Démonstration. Puisque v_i est un vecteur propre associé à la valeur 0 il vient $(1 - z_i)\pi_0 v_i = 0$.

Pour $z_i = 1$ il faut procéder différemment car $(1 - z_i) = 0$. Par (3) on a

$$\begin{aligned} \pi_0 \mathbf{1} &= \lim_{z \rightarrow 1} \frac{\Pi(z) K(z) \mathbf{1}}{1 - z} = \lim_{z \rightarrow 1} \Pi(z) \frac{K(z) \mathbf{1} - K(1) \mathbf{1}}{1 - z} = \lim_{z \rightarrow 1} \Pi(z) \left(1 - \rho^j z \frac{1 - z^{m^j(i)}}{1 - z} \right)_{i=1, \dots, l} \\ &= \Pi(1) (1 - \rho^j m^j(i) z)_{i=1, \dots, l}. \end{aligned}$$

Puisque $\Pi(1)_i$ représente la probabilité d'être dans la lettre i et que celle-ci vaut $1/l$ ceci nous donne

$$\pi_0 \mathbf{1} = \frac{1}{l} \sum_{k=1}^l (1 - \rho^j m^j(k)) = 1 - \frac{a^j \rho^j}{l}.$$

□

La valeur de π_0 ainsi obtenue nous permet de calculer $\Pi(z)$: $\Pi(z) = (1 - z)\pi_0 K(z)^{-1}$. Mais si nous avons calculé une solution de l'équation (2) rien ne nous prouve à ce stade que cette équation a une solution unique et donc que la solution trouvée est la probabilité invariante.

Théorème 7 (Unicité de π). *Le processus X_t est irréductible et récurrent positif et donc π existe et est unique.*

Démonstration. Du théorème 3 nous savons que X est irréductible et récurrent. Comme de plus on peut montrer en utilisant le critère de Foster que la chaîne incluse est récurrente positive et comme l'intensité du processus d'entrée est bornée alors le processus X est irréductible récurrent positif ([Ros83]). Ceci implique ([Ros83]) que la probabilité invariante π existe et est unique. □

Soit $\mathbb{E}(N^j(m^j))$ l'espérance du nombre de clients dans la file \emptyset et $\mathbb{E}(W^j(m^j))$ l'espérance du temps d'attente lorsque le processus de Poisson est filtré par μ . Puisque $\mathbb{E}(N^j(m^j)) = \Pi'(1)\mathbf{1}$ il vient

$$\mathbb{E}(N) = \frac{d}{dz}(\Pi(z)\mathbf{1})|_{z=1} = \pi_0 \frac{d}{dz}((1 - z)\hat{K}(z))|_{z=1}.$$

Sachant que $K(z)$ n'étant pas inversible on introduit la matrice $\hat{K}(z)$ tel que $K(z)\hat{K}(z) = \mathbf{1}$. Par la formule de Little (voir par exemple [BB92]) le temps d'attente est donné par :

$$\mathbb{E}(W^j(m^j)) = \frac{l}{a^j \lambda} \mathbb{E}(N^j(m^j)) - \frac{1}{\mu^j}. \quad (5)$$

De plus par [AGH00b] nous avons que le processus échantillonné par un *shift* du mot \hat{m} a les mêmes performances ainsi $\mathbb{E}(N^j(m^j)) = \mathbb{E}(N^j(S_k(m^j)))$ pour tout $k \in \{1..l\}$.

3.2 Comparaison avec la méthode matricielle géométrique

La méthode “*matrix geometric*” introduite et étudiée dans [Neu81] permet de trouver la probabilité stationnaire de processus de quasi naissance et mort. Afin de calculer la probabilité stationnaire π , on suppose, qu'il existe une matrice R telle que, comme pour la file $M/M/1$, on ait $\pi_{n+1} = \pi_n R$. Cette matrice R appelée matrice de taux obéit à l'équation $RA_1 + R^2 A_2 = 0$, ainsi qu'à une condition de normalisation. Pour la calculer on utilise la méthode itérative suivante : soit $R(n)$ la $i^{\text{ème}}$ itération qui approche R . On pose $R(0) = 0$ et $R(n+1) = -(A_0 + R^2(n)A_2)A_1^{-1}$. Cette solution est coûteuse en temps de calcul même s'il existe des variantes plus efficaces ([LR93]). L'autre principal défaut est que la convergence de $R(n)$ vers R n'est pas assurée.

3.3 Calculs numériques rapides des racines

Lors de nos futurs calculs afin de trouver les politiques optimales nous serons amenés à l'étude de mots dont la longueur pourra être assez importante. C'est pourquoi nous introduisons dès maintenant des moyens visant à améliorer la vitesse des calculs numériques. Pour se faire il convient particulièrement d'accélérer le calcul des racines de l'équation $\Delta(z) = 0$. Pour alléger les notations, il est dans cette partie noté ρ . Le déterminant vaut $\Delta(z) = (-1)^{l+1} \rho^l z^{l+a} + (1 - (1 + \rho)z)^l$. Cette équation $\Delta(z) = 0$ peut donc se réécrire comme suit

$$z^{l+a} = \left(\frac{\rho+1}{\rho}\right)^l \left(z - \frac{1}{\rho+1}\right)^l. \quad (6)$$

En appliquant le théorème de Rouché à l'équation (6) on montre que $\Delta(z)$ est un polynôme qui a l racines dans le disque unité sous la condition $\rho^l/l < 1$ qui est la condition de stabilité de la file.

En posant $x_0 = (\rho + 1)^{-1}$ et en passant au coordonnées polaires i.e. $z = re^{i\theta}$ l'équation (6) devient

$$x_0^2 + 2rx_0 \cos(\theta) + r^2 = \left(\frac{\rho+1}{\rho}\right)^{2l/(l+a)} r^{2l/(l+a)}. \quad (7)$$

Les racines de $\Delta(z)$ sont donc localisées sur la courbe Γ dont l'équation est (7). Cette courbe Γ est un ovale symétrique par rapport à l'axe des x et intersectant celui-ci aux points $(1,0)$ et $(x_0,0)$. Une étude trigonométrique plus poussée de (7) permet d'affirmer que pour chaque r dans l'intervalle $[x_0, \rho/(\rho+1)]$ il existe une solution unique $\theta \in [0, \pi]$ (à laquelle il faut ajouter la valeur symétrique $-\theta$). On définit alors en extrayant le cosinus la fonction $\theta(r)$ par

$$\theta(r) = \arccos \left[\frac{1}{2rx_0} \left(\left(\frac{\rho+1}{\rho}\right)^{2l/(l+a)} r^{2l/(l+a)} - r^2 - x_0^2 \right) \right].$$

Ainsi les racines de $\Delta(z)$ sont les points situés sur Γ tel qu'il existe $i \in \mathbb{Z}$ tel que

$$\arg z^{l+a} = \arg((\rho+1)/\rho)^l (z - x_0)^l + 2i\pi,$$

ce qui donne

$$\theta(r) = \frac{l+a}{l} \arctan \left(\frac{r \sin \theta(r)}{x_0 + r \cos \theta(r)} \right) + \frac{2i\pi}{l}. \quad (8)$$

La recherche des racines de $\Delta(z)$ se décompose alors de la manière suivante :

- 1) Résoudre dans \mathbb{R} l'équation $x^{2(l+a)} = ((\rho+1)/\rho)^l (x - x_0)^{2l}$ ce qui donne x_0 .
- 2) Résoudre (8) pour i variant de 0 à $l/2$. Ceci nous donne la moitié des racines. Les racines conjuguées étant celles situées dans le demi plan des x négatifs.

4 Routage optimal

Nous nous intéressons dans cette section au routage en *boucle ouverte* dans un système dont le processus d'entrée est un processus de Poisson d'intensité λ . Les services quant à eux suivent des lois exponentielles de paramètres μ^1 dans la file Q^1 et μ^2 dans la file Q^2 . Notre but est de trouver la politique de routage optimale, au sens où cette politique va minimiser le temps d'attente moyen des clients dans le système. Dans un routage en boucle ouverte comme le routage ne dépend pas de l'état du système la suite d'allocations optimale m est un mot infini donné a priori et qui va affecter les clients au cours du temps comme décrit dans la section 2. Ainsi notre problème consiste à calculer hors-ligne la suite d'allocation m qui soit optimale.

On définit le temps d'attente moyen d'une séquence d'entrée m par $W(m)$ qui se calcule par

$$W(m) = \lim_{N \rightarrow \infty} \sup \frac{1}{N} \sum_{k=1}^{k=N} W(m)_k$$

où $W(m)_k$ est le temps d'attente du $k^{ième}$ client. En conditionnant par le système choisi la fonction de coût $W(m)$ devient

$$W(m) = \lim_{N \rightarrow \infty} \sup \frac{1}{N} \left(\sum_{k=1}^{k=N} m^1(k) \sum_{k=1}^{k=N} W^1(m^1)_k \right) + \lim_{N \rightarrow \infty} \sup \frac{1}{N} \left(\sum_{k=1}^{k=N} m^2(k) \sum_{k=1}^{k=N} W^2(m^2)_k \right),$$

où $W(m^j)_k$ est le temps d'attente du k^{ieme} client introduit dans la file j . Par le théorème ergodique on obtient

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \left(\sum_{k=1}^{k=N} W^j(m^j)_k \right) = \mathbb{E}(W^j(m^j)).$$

Lorsque le mot m^j est périodique $\mathbb{E}(W^j(m^j))$ est calculée par la méthode du noyau et est ainsi donnée par l'équation (5).

Notre problème peut alors se décomposer en deux étapes. L'une qualitative qui permet de caractériser la forme de la politique optimale et l'autre quantitative qui permet de déterminer les quantités de clients à envoyer dans chaque file.

4.1 Politique de routage optimale

Cette partie est dévolue aux résultats sur la forme de la politique optimale. On introduit $\mathcal{A} = \{0, 1\}$ l'alphabet. Le monoïde libre \mathcal{A}^* est l'ensemble des mots finis issus de \mathcal{A} . Un mot infini est lui un élément de l'ensemble $\mathcal{A}^{\mathbb{N}}$.

Afin de caractériser la politique optimale parmi tous les mots binaires on introduit les mots mécaniques (appelés aussi mot de Sturm). Un mot mécanique est défini par

Définition 8 (Mots mécaniques). *Le mot mécanique supérieur de pente α est le mot infini \overline{m}_α dans lequel la n^{ieme} lettre, pour $n \geq 0$, est :*

$$\overline{m}_\alpha(n) = \lceil (n+1) \times \alpha \rceil - \lceil n \times \alpha \rceil.$$

Le mot mécanique inférieur de pente α est le mot infini \underline{m}_α dans lequel la n^{ieme} lettre, pour $n \geq 0$ est :

$$\underline{m}_\alpha(n) = \lfloor (n+1) \times \alpha \rfloor - \lfloor n \times \alpha \rfloor.$$

Lorsque la pente est rationnelle ($\alpha := a/l$) les mots \overline{m}_α et \underline{m}_α sont périodiques de période l . Si la pente α est irrationnelle les mots mécaniques sont alors apériodiques. Un certain nombre de propriétés des mots mécaniques sont prouvées dans [Lot02, GH01].

La forme de la politique optimale a été étudiée dans [AGH00b, AGH00a].

Théorème 9 ([AGH00b]). *Le temps d'attente moyen est minimisé par une suite d'admission m lorsque celle-ci est un mot mécanique inférieur avec une pente donnée α .*

Démonstration. Ce résultat provient des résultats d'optimisation multicritères obtenus dans [AGH00b] \square

Comme le mot mécanique supérieur est un shift du mot mécanique inférieur et donc que leurs performances sont égales ([AGH00b]) nous nous intéresserons dans la suite aux mots mécaniques supérieurs. Donnons les conséquences de ce théorème sur les mots m^1 et m^2 . Si $m = \overline{m}_\alpha$ alors $m^1 = \overline{m}_\alpha$ et $m^2 = \underline{m}_{1-\alpha}$.

Cependant si la forme de la suite optimale des décisions de routage est connue, aucun moyen n'est donné pour calculer le taux α . C'est l'objet des parties suivantes.

4.2 Taux optimal

Cette partie est dévolue au calcul effectif des taux de clients à envoyer dans chaque file lorsque la politique d'admission est sturmienne. Précisons tout d'abord quels sont les taux de clients qui sont acceptables pour notre problème. Il s'agit des taux pour lesquels le système composé des deux files Q^2 est stable. Pour que le système global soit stable il faut que $\lambda/(\mu + \mu^2) < 1$. Mais chacune des deux files considérée doit aussi être stable. Ceci nous définit alors l'intervalle de stabilité noté \mathcal{I}_s dans lequel α varie. Ainsi par (1) nous avons

$$\mathcal{I}_s = \left] \max\left(0, 1 - \frac{\mu^2}{\lambda}\right), \min\left(1, \frac{\mu^1}{\lambda}\right) \right[.$$

Lorsque la suite d'allocations est \overline{m}_α la fonction de coût à optimiser devient

$$W(\overline{m}_\alpha) = \alpha \mathbb{E}(W^1(\overline{m}_\alpha)) + (1 - \alpha) \mathbb{E}(W^2(\overline{m}_{1-\alpha})). \quad (9)$$

Lorsque α est rationnel $W(\overline{m}_\alpha)$ est déduit de (5).

Notre but est de trouver α_{opt} tel que

$$\alpha_{opt} = \arg \min_{\alpha \in I_s} W(\bar{m}_\alpha).$$

Pour cela nous devons montrer les propriétés suivantes

Théorème 10 (Propriétés de $\mathbb{E}(W^j(\bar{m}_\alpha))$). *La fonction $\alpha \mapsto \alpha \mathbb{E}(W^j(\bar{m}_\alpha))$ est croissante convexe en α et tend vers l'infini quand α tend vers μ^j/λ .*

Démonstration. La croissance s'obtient par des arguments de comparaison de trajectoire (voir [LNT95]). La convexité est issue de [HvdL]. Tandis que la limite est un résultat classique de la théorie des files d'attente. \square

Par conséquent la fonction $\alpha \mapsto \mathbb{E}\alpha(W^j(\bar{m}_\alpha))$ est continue ce qui nous permet de connaître la valeur de l'espérance du temps moyen pour des mots mécaniques infini. De plus par (9) la fonction $\alpha \mapsto W(\bar{m}_\alpha)$ est aussi convexe et tend vers l'infini quand α tend vers les bornes de I_s . Ceci nous prouve par des résultats usuels d'optimisation convexe ([Fle87]) que cette fonction admet un minimum global α_{opt} . Cependant la stricte convexité n'étant pas encore prouvée il n'est pas sûr à l'heure actuelle que ce minimum global soit unique.

5 Applications numériques

5.1 Algorithme du calcul

Nous appliquons ici un algorithme de recherche planifiée qui s'apparente à une dichotomie. Il fonctionne sur le principe suivant : si la fonction à minimiser est convexe soient deux points a et b avec $a < b$ dans un intervalle. Si $W(\bar{m}_a) \leq W(\bar{m}_b)$ alors $\alpha_{opt} \leq b$ par contre si $W(\bar{m}_a) \geq W(\bar{m}_b)$ alors $\alpha_{opt} \geq a$. Ceci nous permet de diminuer l'intervalle de recherche à chaque pas. Cependant la fonction de coût pour être calculée nécessite que les points utilisés dans la recherche soient rationnels. De plus le temps de calcul étant fortement lié à la longueur du mot nous privilégions la recherche de rationnels dont le dénominateur est le plus petit possible. Ceci est fait en utilisant le fait que si nous avons deux rationnels p_1/q_1 et p_2/q_2 alors le rationnel de plus petit dénominateur dans $]p/q_1, p_2/q_2[$ est $p_1 + p_2/q_1 + q_2$.

5.2 Calcul effectif des taux

Nous fixons ici $\mu^1 = 7/16$ et $\mu^2 = 3\mu^1 = 21/16$. Nous faisons alors varier la charge totale du système entre 0 et 1 exclu en faisant varier les inter arrivées. L'intensité λ du processus de Poisson varie ainsi entre 0 et 28/16. Sur la figure 1 on a représenté le taux optimal (calculé avec une précision de 10^{-3}) en fonction de la charge. Nous avons aussi représenté sur la figure 1 le taux $\alpha = 1/4$ par une ligne en pointillés. Ce taux donne une politique qui consiste à envoyer 3 fois plus de clients dans la file la plus rapide \mathcal{Q} . Comme on peut le voir le taux $\alpha_{opt} = 1/4$ n'est pas toujours optimal.

Références

- [AGH00a] E. Altman, B. Gaujal, and A. Hordijk. Admission control in stochastic event graphs. *IEEE Trans. Aut. Cont.*, 45(5):854–867, May 2000.
- [AGH00b] E. Altman, B. Gaujal, and A. Hordijk. Balanced sequences and optimal routing. *J. Assoc. Comput. Mach.*, 47:752–775, 2000.
- [Alt00] E. Altman. Applications of markov decision processes in communication networks : a survey. Technical Report RR-3984, INRIA, Aout 2000.
- [BB92] F. Baccelli and P. Bremaud. *Elements of queueing theory*. Springer, 1992.
- [Fle87] R. Fletcher. *Practical methods of optimization*. Wiley, 1987.
- [GH01] B. Gaujal and E. Hyon. Optimal routing policies in two deterministic queues. *Réseaux et systèmes répartis - Calculateurs Parallèles*, 13(6):601–633, 2001.
- [GH02] B. Gaujal and E. Hyon. Optimal routing in deterministic queues in tandem. Technical Report RR-4393, INRIA, March 2002.

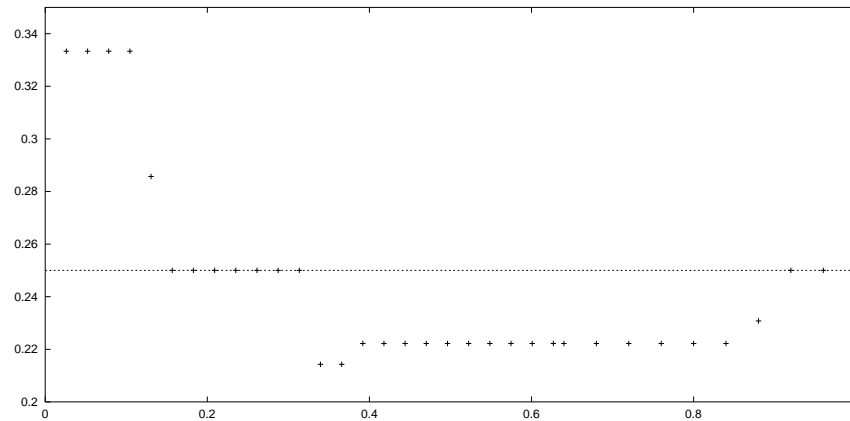


FIG. 1 – Taux optimaux

- [HvdL] A. Hordijk and D. van der Laan. On the convexity of the stationary waiting time as a function of the routing densities. Private Communication.
- [JM97] A. Jean-Marie. The waiting time distribution in poisson-driven deterministic systems. Technical Report RR-3083, INRIA, 1997.
- [LNT95] Z. Liu, P. Nain, and D. Towsley. Sample path methods in the control of queues. *Queueing Systems*, (21):293–335, 1995.
- [Lot02] M. Lothaire. *Algebraic Combinatorics on Words*, chapter Sturmian Words. Cambridge University Press, 2002.
- [LR93] G. Latouche and V. Ramaswami. A logarithmic reduction algorithm for quasi-birth-death processes. *Journal of Applied Probability*, 30:650–674, 1993.
- [Neu81] M.F. Neuts. *Matrix-Geometric Solutions in stochastic Models An Algorithmic Approach*. John Hopkins University Press, 1981.
- [Ros83] S.M. Ross. *Stochastic Processes*. Wiley, 1983.
- [Vil99] C. Villamizar. Mpls optimized multipath (mpls-omp). Internet draft (work in progress), IETF, <ftp://ftp.isi.edu/internet-drafts/draft-villamizar-mpls-omp-01>, 1999.